



OPEN ACCESS

EDITED BY

Philipp Beckerle,
University of Erlangen-Nuremberg, Germany

REVIEWED BY

Merle Theresa Fairhurst,
Dresden University of Technology, Germany
Ioannis Delis,
University of Leeds, United Kingdom

*CORRESPONDENCE

Mattia Pinaridi
✉ m.pinaridi@unicampus.it

RECEIVED 20 March 2023

ACCEPTED 30 May 2023

PUBLISHED 15 June 2023

CITATION

Pinaridi M, Di Stefano N, Di Pino G and
Spence C (2023) Exploring crossmodal
correspondences for future research
in human movement augmentation.
Front. Psychol. 14:1190103.
doi: 10.3389/fpsyg.2023.1190103

COPYRIGHT

© 2023 Pinaridi, Di Stefano, Di Pino and
Spence. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Exploring crossmodal correspondences for future research in human movement augmentation

Mattia Pinaridi^{1*}, Nicola Di Stefano², Giovanni Di Pino¹ and Charles Spence³

¹NeXT Lab, Neurophysiology and Neuroengineering of Human-Technology Interaction Research Unit, Università Campus Bio-Medico di Roma, Rome, Italy, ²Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy, ³Crossmodal Research Laboratory, University of Oxford, Oxford, United Kingdom

“Crossmodal correspondences” are the consistent mappings between perceptual dimensions or stimuli from different sensory domains, which have been widely observed in the general population and investigated by experimental psychologists in recent years. At the same time, the emerging field of human movement augmentation (i.e., the enhancement of an individual’s motor abilities by means of artificial devices) has been struggling with the question of how to relay supplementary information concerning the state of the artificial device and its interaction with the environment to the user, which may help the latter to control the device more effectively. To date, this challenge has not been explicitly addressed by capitalizing on our emerging knowledge concerning crossmodal correspondences, despite these being tightly related to multisensory integration. In this perspective paper, we introduce some of the latest research findings on the crossmodal correspondences and their potential role in human augmentation. We then consider three ways in which the former might impact the latter, and the feasibility of this process. First, crossmodal correspondences, given the documented effect on attentional processing, might facilitate the integration of device status information (e.g., concerning position) coming from different sensory modalities (e.g., haptic and visual), thus increasing their usefulness for motor control and embodiment. Second, by capitalizing on their widespread and seemingly spontaneous nature, crossmodal correspondences might be exploited to reduce the cognitive burden caused by additional sensory inputs and the time required for the human brain to adapt the representation of the body to the presence of the artificial device. Third, to accomplish the first two points, the benefits of crossmodal correspondences should be maintained even after sensory substitution, a strategy commonly used when implementing supplementary feedback.

KEYWORDS

crossmodal correspondence, augmentation, multisensory integration, sensory feedback, embodiment

1. Introduction

1.1. Crossmodal correspondences

“Crossmodal correspondences” are the consistent matchings between perceptual dimensions or stimulus attributes from different sensory domains that are observed in normal observers (i.e., in non-synesthetes; see [Spence, 2011](#), for a review). One of the most famous audiovisual associations was discovered almost a century ago by [Köhler \(1929\)](#). In particular, the German psychologist observed that people tend to associate the term “baluba” with curved lines, while the term “takete” is associated with angular shapes instead. To date, the existence of a very wide range of such crossmodal associations have been demonstrated, involving most, if not all, combinations of senses (see [Spence, 2011](#), for a review). Associations involving auditory and visual stimuli have been predominantly addressed. That said, combinations involving haptics ([Slobodenyuk et al., 2015](#); [Wright et al., 2017](#); [Lin et al., 2021](#)) and/or the chemical senses (i.e., taste and gustation) have increasingly been investigated over the last decade, also in relation to the growing interest in research on food and wine pairing, and food marketing/packaging design ([North, 2012](#); [Spence and Wang, 2015](#); [Pramudya et al., 2020](#); [Spence, 2020](#)).

Crossmodal associations have primarily been investigated in the field of experimental psychology/perception research, in which it has, on occasion, been related to the wider phenomenon of multisensory integration ([Parise and Spence, 2009](#)), conceived of as the processing and organization of sensory inputs received from different pathways into a unified whole ([Tong et al., 2020](#)). A number of intriguing questions concerning the nature of crossmodal correspondences and the relative contribution of different factors remains open, with some researchers putting forward the existence of intersensory, or suprasensory stimulus qualities, such as brightness, intensity, roughness, that can be picked-up by multiple senses (e.g., see, [von Hornbostel, 1931](#); [Walker-Andrews et al., 1994](#)), and others focusing more on the integration of information occurring in the brain, as a result of crossmodal binding mechanisms ([Calvert and Thesen, 2004](#)).

1.2. Multisensory processing and feedback in the context of human augmentation

The mechanisms behind the processing of multisensory stimuli are also of great interest to scientists seeking feedback strategies to be implemented in the context of human motor control. Indeed, multisensory feedback has been shown to improve an individual's ability to command their body significantly, especially in complex interactions requiring constant adjustment and recalibration ([Wolpert et al., 1995](#); [Diedrichsen et al., 2010](#); [Shadmehr et al., 2010](#)). The same benefit can potentially also be extended to the control of artificial devices, such as teleoperated robots or prostheses. In the field of prosthetics research, researchers have endowed prostheses with artificial senses, so that users can retrieve sensory information from these artificial devices as they would with their natural limbs, ultimately leading to better control and higher acceptability ([Di Pino et al., 2009](#); [Di Pino et al., 2012, 2020](#);

[Raspopovic et al., 2014](#); [Marasco et al., 2018](#); [Page et al., 2018](#); [Zollo et al., 2019](#)).

In recent years, the role of multisensory feedback has been studied within the framework of human augmentation, an emerging field that aims at enhancing human abilities beyond the level that is typically attainable by able-bodied users ([Di Pino et al., 2014](#); [Eden et al., 2021](#)). Movement augmentation can, for instance, be achieved through the use of supernumerary robotic limbs (SRLs), artificial devices that can be controlled simultaneously with, but independently from, the natural limbs, thus opening up new possibilities in the interaction with the environment. SRLs can be controlled through the movement of body parts that may not be directly involved in the task, actuated by means of joysticks, trackers, or retrieved from the electrical muscle activity ([Eden et al., 2021](#)). It turns out that the quality of the motor interface is of pivotal importance as far as achieving proficient performance is concerned. However, proficient control can only be achieved by closing the sensorimotor loop between user and robotic device by means of reliable feedback, possibly covering multiple senses. Indeed, it has recently been demonstrated how supplementary sensory feedback improved the regulation of the force exerted by the SRLs end-effector when interacting with an object ([Hussain et al., 2015](#); [Hernandez et al., 2019](#); [Guggenheim and Asada, 2021](#)), as well as increasing the accuracy in reaching a target or replicating the end-effector position ([Segura Meraz et al., 2017](#); [Aoyama et al., 2019](#); [Pinardi et al., 2021, 2023](#)). It can also reduce the amount of time required to complete behavioral tasks ([Hussain et al., 2015](#); [Sobajima et al., 2016](#); [Noccaro et al., 2020](#)).

Multisensory integration plays a key role in building body representations, the map that our brain uses to recognize and identify our body and the model the brain uses to control its movement (i.e., body image and body schema; [de Vignemont, 2010](#); [Blanke, 2012](#); [Moseley et al., 2012](#)). Sensory feedback continuously updates these representations, allowing the brain to accept artificial limbs, as has frequently been demonstrated in the famous Rubber Hand Illusion ([Botvinick and Cohen, 1998](#)), where a visible rubber hand and the participant's own hidden hand are brushed synchronously (though see [Pavani et al., 2000](#), for evidence that synchronous stroking is not always required to induce the illusion). As a result of this visuotactile congruency, the artificial limb is included in the body representation and thus perceived as belonging to the body of the person experiencing the illusion. This process, often labeled “embodiment” ([Botvinick and Cohen, 1998](#); [Pinardi et al., 2020a](#)), has been reported for several types of artificial limbs, including prostheses, and has been linked to an improvement in artificial limb control and acceptability ([Marasco et al., 2011](#); [Gouzien et al., 2017](#)). Recent works suggests that embodiment might be perceived for SRLs as well, making it a relevant topic in the field of human augmentation ([Di Pino et al., 2014](#); [Arai et al., 2022](#); [Di Stefano et al., 2022](#); [Umezawa et al., 2022](#)).

1.3. Perspectives for future research in human augmentation

In this paper, we present three novel considerations concerning how future research on human augmentation might benefit from the vast knowledge concerning crossmodal correspondences. First,

crossmodal correspondences could make multisensory feedback from the augmenting device easier to process and understand. Second, the spontaneous and widespread mechanisms that characterize crossmodal correspondences might well be expected to reduce the amount of time that is needed for users to learn and decode the supplementary feedback of SRLs and to embody it faster. Third, the benefits of crossmodal correspondences involving two sensory modalities should persist if one modality is transduced but the informative content is maintained (see [Figure 1](#)).

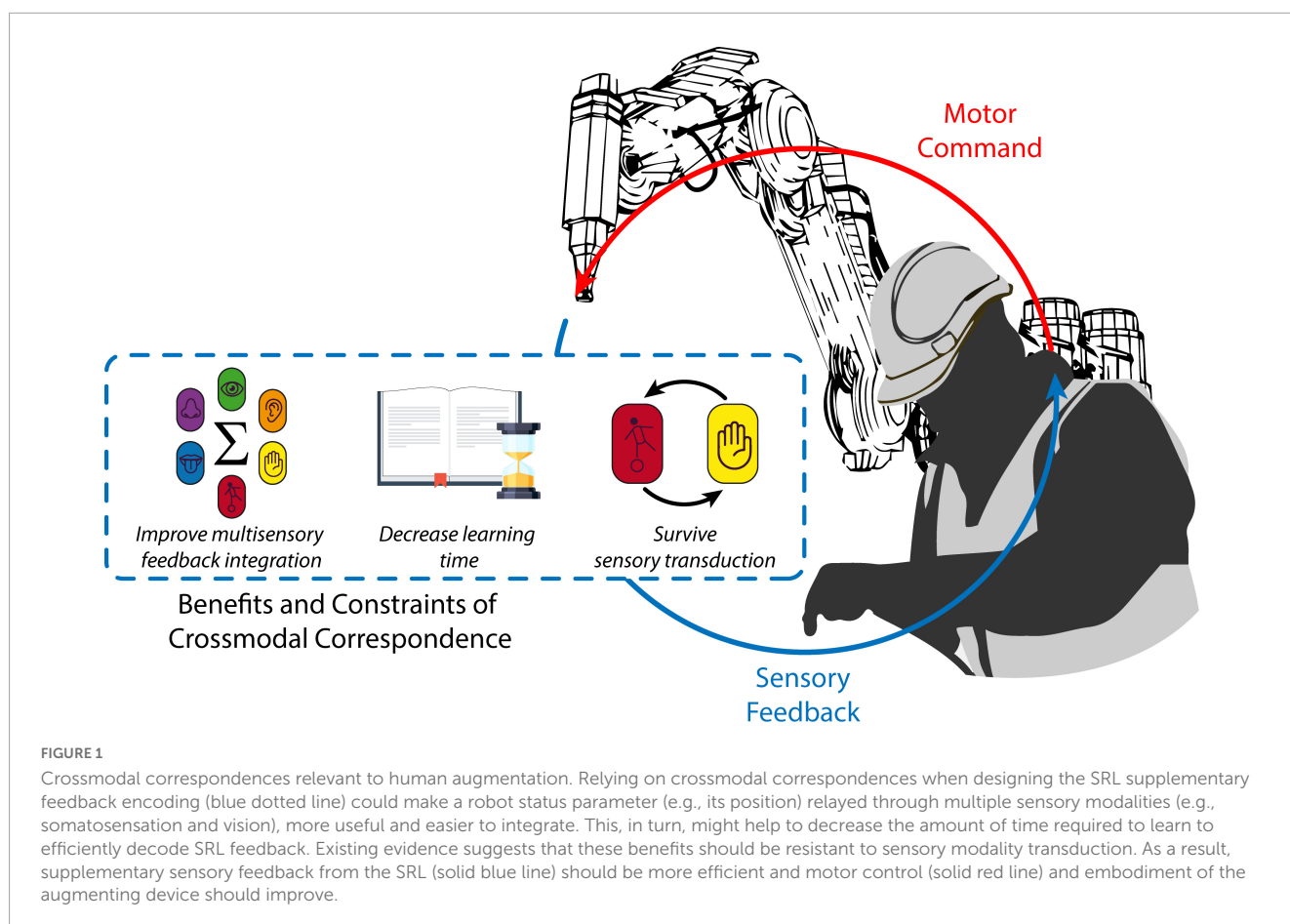
2. Improving supplementary feedback integration

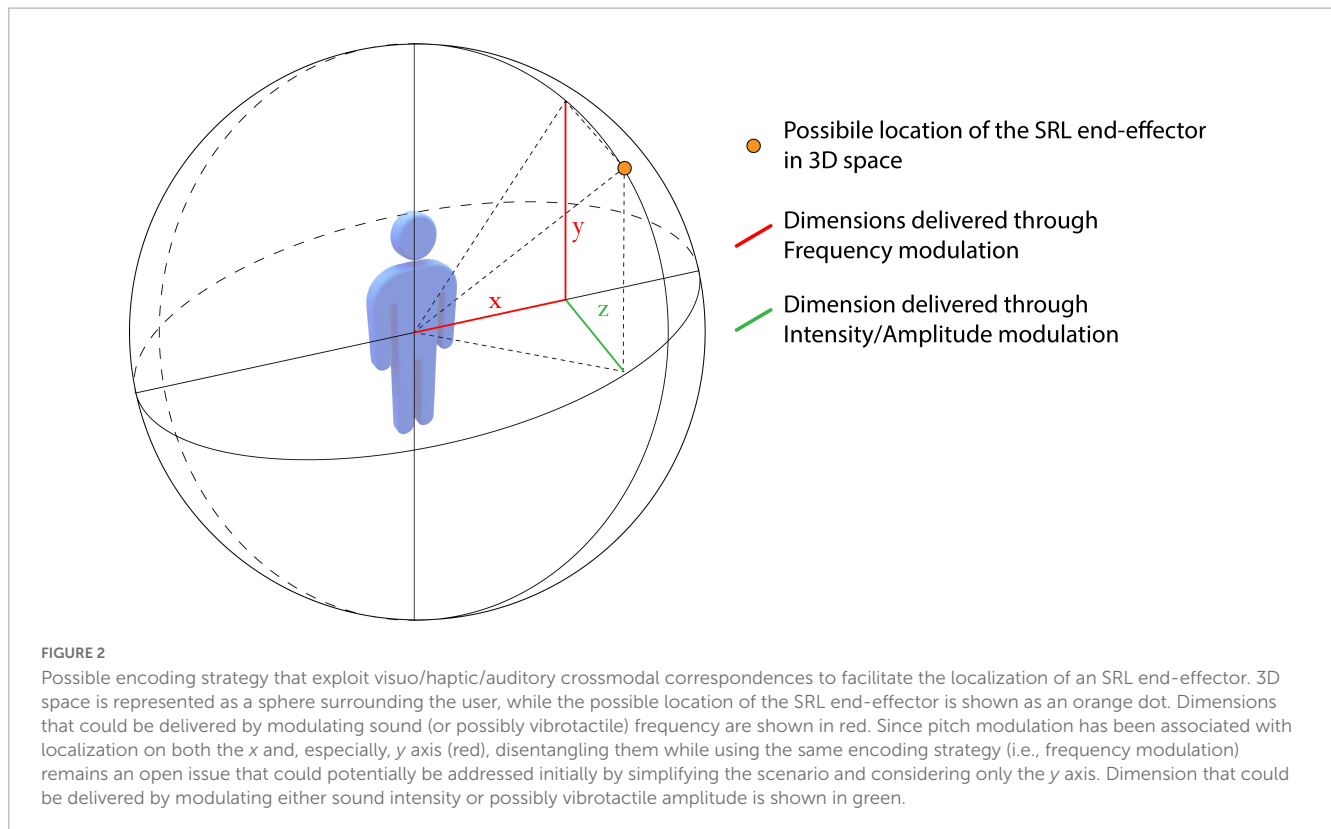
Crossmodal correspondences modulate multisensory integration (see [Spence, 2011](#), for a review). More specifically, crossmodally congruent multisensory stimuli possess a perceptual bond that leads to increased strength/likelihood of integration thus potentially reducing spatiotemporal discrepancy. Indeed, when participants had to judge the temporal order of auditory and visual stimuli presented in pairs, worse performance was observed with pairs of stimuli that were crossmodally congruent as compared to those that were incongruent ([Parise and Spence, 2009](#)).

Crossmodally congruent pairs of audiovisual stimuli give rise to larger spatial ventriloquism effects (i.e., mislocalization of the source of a sound in space toward an irrelevant visual stimulus) compared to incongruent pairs ([Parise and Spence,](#)

[2008](#)). Crossmodal correspondences can also impact the perceived direction of moving visual stimuli: participants perceived gratings with ambiguous motion as moving upward when coupled with an ascending pitch sound and vice versa ([Maeda et al., 2004](#)).

Additional research findings have also demonstrated that crossmodal correspondences can improve behavioral performance in sensorimotor tasks. For instance, in a speeded target discrimination task in which participants were asked whether the pairs of color and sound stimuli that were presented appropriately matched, higher accuracy and faster reaction times were obtained with crossmodally congruent pairs of stimuli as compared to incongruent ones ([Marks, 1987](#); [Klapetek et al., 2012](#); [Sun et al., 2018](#)). Recently, audiovisual correspondences were shown to promote higher accuracy and faster responses during a reaction time task when stimuli were congruent rather than incongruent ([Ihalainen et al., 2023](#)). Even complex motor tasks, such as drawing, are affected by crossmodal correspondences, with participants obtaining higher accuracy in representing graphically certain features of auditory stimuli (e.g., pitch) compared to others (e.g., loudness) ([Küssner and Leech-Wilkinson, 2014](#)). Therefore, a first suggestion is that crossmodal correspondences can be exploited to improve the integration of multisensory information concerning the status of the SRL. According to the Bayesian integration framework ([Körding and Wolpert, 2004](#); [Chandrasekaran, 2017](#)), supplementary feedback is particularly useful when it carries information with lower “noise” (i.e., less disturbance in





the signal) compared to other sensory cues. Hence, if the user is operating the SRL in an environment with reduced visibility (e.g., presence of smoke or dust, or darkness for workers operating under water), supplementary feedback carrying information on the motion of the robot would become particularly useful. This usefulness could be increased even more by capitalizing on the correspondence between perceived motion and acoustic pitch. For instance, playing an ascending pitch when the robot moves upward would easily bias the attention of the user toward that direction, because of the integration guided by crossmodal correspondence, thus improving the interaction under low visibility conditions.

3. Reducing learning time

A striking feature of crossmodal correspondences is how spontaneous and widespread they appear to be. Whether this perceptual phenomenon reflects an innate and universal mechanism is not yet clear, since research has produced conflicting evidence (Ludwig et al., 2011; Spence and Deroy, 2012). Despite an experience-based explanation for certain crossmodal correspondences having been proposed (Speed et al., 2021), and possible cultural differences between groups occasionally being reported (Spence, 2022), the empirical evidence that has been published to date shows that most healthy adults perceive crossmodal correspondence spontaneously, consistently, and without the need for any specific training (though see also Klapetek et al., 2012). The possibility of associating information from different sensory modalities without a dedicated training is of great interest for the field of human augmentation.

Indeed, learning to use a sensory augmentation tool requires the user to become familiar with its physical features (e.g., dimensions, weight, joint stiffness, and degrees of freedom) in order to interact with it proficiently. Rich supplementary feedback can greatly facilitate this process, but learning to understand and use such feedback often requires extensive training, especially when supplementary feedback carries information that is not available to the other senses. This sensory substitution requires being exposed to both the original sensory feedback (e.g., vision) and the supplementary one (e.g., vibrotactile stimulation) in order to establish an association that is sufficiently strong to survive the removal of the former and thus allow the latter to provide useful information. The latest research demonstrates that this is possible in the specific context of human augmentation (Pinardi et al., 2021; Umezawa et al., 2022). However the long training sessions required to fully benefit from supplementary feedback, can all too easily result in participant exhaustion and demotivation, thus limiting the feasibility of the experimental approach. To avoid this, supplementary feedback usually relies on a simple encoding strategy that can be learned quickly. For instance, it is useful to respect the spatial distribution between workspace organization and feedback interface on the participant's body (i.e., feedback of proximal joints is delivered proximally on the participant's body; Noccaro et al., 2020; Pinardi et al., 2021), or to couple a vibrotactile stimulation frequency that is higher with a higher force exerted by the robotic limb (Hussain et al., 2015). By removing the need for feedback decoding, the effort to facilitate feedback learning could be pushed toward a more efficient application by exploiting the spontaneous associations that characterize crossmodal correspondences, such as the well-documented space-pitch/loudness correspondence.

3.1. Application scenario: exploiting space-pitch/loudness correspondence for SRL localization

Pitch is consistently associated with spatial locations on the vertical (Bordeau et al., 2023), and to a lesser extent horizontal, axis, such that a low-pitched sound is more likely to be associated with a left/low location whereas a high pitch is associated with a right/high location (Mudd, 1963; Rusconi et al., 2006; Spence, 2011; Guilbert, 2020). At the same time, however, loudness is known to be an effective indicator of distance, such that, when presented with two sounds which differ only in loudness, people tend to associate the louder sound with a closer sound source, and *vice versa* (Eitan, 2013; Di Stefano, 2022).

Sonification is a strategy of heteromodal sensory substitution which relays kinematic and dynamic features of a movement with sounds (Castro et al., 2021a,b); however, its effectiveness is dependent on the specific rules used for encoding. Sonification can also be exploited to relay feedback related to an SRL. In light of enhancing the intuitiveness and hence usability of sensory substitution devices by means of the incorporation of crossmodal correspondence, we propose to use sound frequency and intensity (which determines perceived pitch and loudness, respectively) to help users to more easily localize the SRL in tri-dimensional space with minimal need for dedicated training, since pitch/space correspondence has been shown to occur spontaneously (Guilbert, 2020), even in new-borns (Walker et al., 2018), and space/loudness is a reliable and consistent association which reflects the regularities in the environment (Eitan, 2013) (see Figure 2). Given that the spatial resolution of pitch/space correspondences has yet to be determined, it is hard to provide clear indications on the effective contribution of this association to the spatial localization of SRLs. However, a rough indication on the region of space occupied by the SRL in terms of the two principal axes (i.e., up/down, and right/left) would be probably enough informative for these kinds of tasks, and we expect that such information can be likely provided intuitively, thus potentially reducing the learning time.

The association between pitch and the vertical/horizontal axes could also be used in human augmentation to generate more “instinctive” warning signals that might direct the operators’ spatial attention in an intuitive way, i.e., with no need for dedicated training (Ho and Spence, 2008). Finally, sonification has been shown to improve the behavioral performance of sportsmen (e.g., Maes et al., 2016; Lorenzoni et al., 2019; O’Brien et al., 2020), and thus coupling it with crossmodal correspondence mechanisms might further facilitate this beneficial effect. Since pitch is associated at the same time with both horizontal and vertical direction, it would be important to disentangle the modulation of these axis. This could be addressed initially by simply considering one of the two axes (2D workspace), and later by acting on how feedback is delivered (e.g., deliver two pitches sequentially, the first coding the x axis and the second coding the y axis). Additionally, while the loudness/distance association has been investigated primarily on the z axis, it should be tested whether it also consistently present on the others.

This perspective is further corroborated by studies involving blind participants, that show that the use of sensory substitution

devices, despite not being devoid of challenges (Barilari et al., 2018; Hamilton-Fletcher et al., 2018; Sourav et al., 2019), is somewhat easier to learn if they are based on crossmodal correspondences (Hamilton-Fletcher et al., 2016) and a similar correspondence-based facilitation has also emerged in the case of language learning as well (Imai et al., 2008, 2015).

Moreover, it can be predicted that the haptic feedback might be used similarly to pitch, assuming that frequency and amplitude of a vibrotactile stimulation are associated with vertical/horizontal localization and with distance, or depth, respectively. Should this be confirmed empirically, it would represent a useful principle to guide the design of supplementary feedback. Additionally, given the multifaceted nature of the somatosensory modality, haptic feedback could potentially convey several informative contents simultaneously, so that multiple stimulus dimensions can be captured more clearly compared to information conveyed through auditory feedback. Indeed, somatosensation includes a great variety of sensory stimuli, such as, vibration, pressure, and skin stretch. Combining different types of stimuli, such as high frequency vibrotactile stimulation and skin stretch, to encode different information (e.g., coordinates on different axis), could reduce the risk of ambiguous encoding while maintaining a high informative value. Though, that being said, one should never forget the inherent limitations associated with the skin as a means of information transfer (see Spence, 2014).

4. Sensory transduction

Studies on human motor control have proven that proprioception is an unobtrusive yet rich sensory modality allowing people to complete dextrous motor tasks with minimal attentional effort (Proske and Gandevia, 2012). This is further confirmed in patients who lost proprioception, and must continuously rely on visual monitoring to carry out daily life actions, with a huge attentional effort and lower learning performance (Danna and Velay, 2017; Miall et al., 2019, 2021). Hence, relaying proprioceptive-like information concerning SRLs (i.e., joint angle, end effector acceleration or exerted force) to the user can result in improved SRL control. However, in the case of supplementary feedback, it is often difficult to adopt a homomodal coding approach (information related to a given sensory modality are fed back to the user exploiting the same sensory modality e.g., pressure on the robot surface with tactors upon the user skin). This is due to the fact that SRLs proprioceptive parameters should be delivered using proprioceptive stimulation, and this is sometimes impossible or impractical (e.g., delivering joint angles by means of a kinesthetic illusion (Pinardi et al., 2020b) imposes heavy constraints on experimental protocols). Hence, translating a proprioceptive parameter of the robot (e.g., joint angle) into a pattern of tactile stimulation (e.g., vibratory stimulation) is easier to implement, resulting in heteromodal feedback. This approach solves a number of technical issues but introduces the need to translate supplementary sensory feedback from one modality (e.g., proprioception) to another (e.g., touch). Hence, the possibility of maintaining the benefits of crossmodal correspondence across different sensory modalities would be an empirical direction worth investigating. For instance, higher auditory pitch is associated with

faster motion (Zhang et al., 2021) and higher physical acceleration (Eitan and Granot, 2006). However, in a human augmentation paradigm, participants would not be directly informed about speed or acceleration through supplementary feedback, but rather with vibratory stimulation that codes those parameters (e.g., higher vibrotactile frequency codes for higher acceleration). Hence, would the association between acceleration (or speed) and auditory pitch still be valid, if acceleration were to be substituted with a vibrotactile stimulation coding the same information?

To the best of our knowledge, there are no studies that directly address this question, however, considering that crossmodal correspondences have been demonstrated between every pair of sensory modalities (North, 2012; Nava et al., 2016; Wright et al., 2017; Speed et al., 2021) and are, by definition, cross-modal, it is likely that the underlying mechanisms would not be disrupted by sensory modality transduction.

5. Discussion and conclusion

In the present perspective paper, possible intersections between two different and, to the best of our knowledge, previously unrelated research fields, namely human augmentation and crossmodal correspondences have been explored. We propose that future research on the control and embodiment of SRLs could benefit from exploiting the spontaneous and widespread mechanisms of crossmodal correspondences, which have been shown to impact on spatiotemporal features and constraints of multisensory integration. The potential improvement of the control of SRLs would likely result as a by-product of the well-documented positive effect of crossmodal correspondence on attentional processing (Spence, 2011; Klapetek et al., 2012; Orchard-Mills et al., 2013; Chiou and Rich, 2015). Indeed, attention is considered a crucial resource when it comes to learning new motor skills and improving motor performance, especially in complex environments or during the performance of difficult tasks (Song, 2019), such as controlling an SRL. The link between motor control, attention and sensory processing is further demonstrated by the fact that multisensory contingencies can modulate the coupling between attention and motor planning (Dignath et al., 2019). However, further evidence is required to assess and disentangle the specific impact of crossmodal correspondences on motor control or attentional processing in the framework of human augmentation.

While the insights suggested here could be extended to any robotic teleoperated device, they are particularly fit for the field of human augmentation. Indeed, SRL users are supposed to control robotic limbs and natural ones at the same, while receiving feedback from both. This produces an additional challenge that is not present in prosthetics or when dealing with robotic teleoperation. The added sensorimotor processing required in human augmentation, and the absence of residual neural resources that can be repurposed for conveying feedback (Makin et al., 2017; Dominijanni et al., 2021) justifies the need for new strategies that minimize the effort required for sensory feedback learning and maximize its usefulness, and even though benefits of crossmodal correspondence have usually been observed in simple tasks (i.e., speeded target discrimination) and control and embodiment of an augmenting

device are much more complex tasks, they ultimately rely on the same mechanisms of multisensory integration.

Finally, we argue that crossmodal correspondences could be exploited to deliver even highly sophisticated feedback with reduced cognitive burden. For instance, by capitalizing on the correspondence between tactile and auditory roughness (see Di Stefano and Spence, 2022, for a review), it might be possible to deliver the sensation of a rough texture through a rough sound (e.g., equal intensity for all frequencies, such as white noise, Pellegrino et al., 2022). More specifically, simply using a camera located on the tip of the end-effector, it would be possible to analyze surface texture, determine its roughness level through a machine-learning algorithm, and modulate the features of an auditory cue to provide information to the user on the surface tactile quality (Guest et al., 2002).

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

MP and NDS conceived the work and wrote the first draft of the manuscript. CS and GDP revised the entire manuscript and provided the insightful comments on all sections. All authors read and approved the submitted version.

Funding

This work was supported by the Italian Ministry of Education, University and Research under the “FARE: Framework Attrazione e Rafforzamento Eccellenze Ricerca in Italia” research program (ENABLE, no. R16ZBLF9E3) and by the European Commission under the “NIMA: Non-invasive Interface for Movement Augmentation” project (H2020-FETOPEN-2018-2020, ID: 899626).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Aoyama, T., Shikida, H., Schatz, R., and Hasegawa, Y. (2019). Operational learning with sensory feedback for controlling a robotic thumb using the posterior auricular muscle. *Adv. Robot.* 33, 243–253. doi: 10.1080/01691864.2019.1566090
- Arai, K., Saito, H., Fukuoka, M., Ueda, S., Sugimoto, M., Kitazaki, M., et al. (2022). Embodiment of supernumerary robotic limbs in virtual reality. *Sci. Rep.* 12:9769. doi: 10.1038/s41598-022-13981-w
- Barilari, M., de Heering, A., Crollen, V., Collignon, O., and Bottini, R. (2018). Is red heavier than yellow even for blind? *i-Perception* 9:2041669518759123. doi: 10.1177/2041669518759123
- Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nat. Rev. Neurosci.* 13, 556–571. doi: 10.1038/nrn3292
- Bordeau, C., Scalvini, F., Migniot, C., Dubois, J., and Ambard, M. (2023). Cross-modal correspondence enhances elevation localization in visual-to-auditory sensory substitution. *Front. Psychol.* 14:1079998. doi: 10.3389/fpsyg.2023.1079998
- Botvinick, M., and Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature* 391:756. doi: 10.1038/35784
- Calvert, G. A., and Thesen, T. (2004). Multisensory integration: Methodological approaches and emerging principles in the human brain. *J. Physiol. Paris* 98, 191–205. doi: 10.1016/j.jphysparis.2004.03.018
- Castro, F., Bryjka, P. A., Di Pino, G., Vuckovic, A., Nowicky, A., and Bishop, D. (2021a). Sonification of combined action observation and motor imagery: Effects on corticospinal excitability. *Brain Cogn.* 152:105768. doi: 10.1016/j.bandc.2021.105768
- Castro, F., Osman, L., Di Pino, G., Vuckovic, A., Nowicky, A., and Bishop, D. (2021b). Does sonification of action simulation training impact corticospinal excitability and audiomotor plasticity? *Exp. Brain Res.* 239, 1489–1505. doi: 10.1007/s00221-021-06069-w
- Chandrasekaran, C. (2017). Computational principles and models of multisensory integration. *Curr. Opin. Neurobiol.* 43, 25–34. doi: 10.1016/j.comb.2016.11.002
- Chiou, R., and Rich, A. N. (2015). Volitional mechanisms mediate the cuing effect of pitch on attention orienting: The influences of perceptual difficulty and response pressure. *Perception* 44, 169–182. doi: 10.1068/p7699
- Danna, J., and Velay, J. L. (2017). On the auditory-proprioception substitution hypothesis: Movement sonification in two deafferented subjects learning to write new characters. *Front. Neurosci.* 11:137. doi: 10.3389/fnins.2017.00137
- de Vignemont, F. (2010). Body schema and body image—pros and cons. *Neuropsychologia* 48, 669–680. doi: 10.1016/j.neuropsychologia.2009.09.022
- Di Pino, G., Guglielmelli, E., and Rossini, P. M. (2009). Neuroplasticity in amputees: Main implications on bidirectional interfacing of cybernetic hand prostheses. *Prog. Neurobiol.* 88, 114–126. doi: 10.1016/j.pneurobio.2009.03.001
- Di Pino, G., Maravita, A., Zollo, L., Guglielmelli, E., and Di Lazzaro, V. (2014). Augmentation-related brain plasticity. *Front. Syst. Neurosci.* 8:109. doi: 10.3389/fnsys.2014.00109
- Di Pino, G., Porcaro, C., Tombini, M., Assenza, G., Pellegrino, G., Tecchio, F., et al. (2012). A neurally-interfaced hand prosthesis tuned inter-hemispheric communication. *Restor. Neurol. Neurosci.* 30, 407–418. doi: 10.3233/RNN-2012-120224
- Di Pino, G., Romano, D., Spaccasassi, C., Mioli, A., D'Alonzo, M., Sacchetti, R., et al. (2020). Sensory- and action-oriented embodiment of neurally-interfaced robotic hand prostheses. *Front. Neurosci.* 14:389. doi: 10.3389/fnins.2020.00389
- Di Stefano, N. (2022). The spatiality of sounds. From sound-source localization to musical spaces. *Aisthesis. Pratiche, linguaggi e saperi dell'estetico* 15, 173–185.
- Di Stefano, N., and Spence, C. (2022). Roughness perception: A multisensory/crossmodal perspective. *Atten. Percept. Psychophys.* 84, 2087–2114. doi: 10.3758/s13414-022-02550-y
- Di Stefano, N., Jarrassé, N., and Valera, L. (2022). The ethics of supernumerary robotic limbs. An enactivist approach. *Sci. Eng. Ethics* 28:57. doi: 10.1007/s11948-022-00405-1
- Diedrichsen, J., Shadmehr, R., and Ivry, R. B. (2010). The coordination of movement: Optimal feedback control and beyond. *Trends Cogn. Sci.* 14, 31–39. doi: 10.1016/j.tics.2009.11.004
- Dignath, D., Herbort, O., Pieczykolan, A., Huestegge, L., and Kiesel, A. (2019). Flexible coupling of covert spatial attention and motor planning based on learned spatial contingencies. *Psychol. Res.* 83, 476–484. doi: 10.1007/s00426-018-1134-0
- Dominijanni, G., Shokur, G., Salvietti, G., Buehler, S., Palmerini, E., Rossi, S., et al. (2021). Enhancing human bodies with extra robotic arms and fingers: The neural resource allocation problem. *Nat. Mach. Intell.* 3, 850–860.
- Eden, J., Bräcklein, M., Pereda, J. I., Barsakcioglu, D. Y., Di Pino, G., Farina, D., et al. (2021). Human movement augmentation and how to make it a reality. *arXiv preprint. arXiv:2106.08129*.
- Eitan, Z. (2013). "How pitch and loudness shape musical space and motion," in *The psychology of music in multimedia*, eds S.-L. Tan, A. J. Cohen, S. D. Lipscomb, and R. A. Kendall (Oxford: Oxford University Press), 165–191.
- Eitan, Z., and Granot, R. Y. (2006). How music moves: Musical parameters and listeners images of motion. *Music Percept.* 23, 221–248.
- Gouzien, A., de Vignemont, F., Touillet, A., Martinet, N., De Graaf, J., Jarrassé, N., et al. (2017). Reachability and the sense of embodiment in amputees using prostheses. *Sci. Rep.* 7:4999. doi: 10.1038/s41598-017-05094-6
- Guest, S., Catmur, C., Lloyd, D., and Spence, C. (2002). Audiotactile interactions in roughness perception. *Exp. Brain Res.* 146, 161–171. doi: 10.1007/s00221-002-1164-z
- Guggenheim, J. W., and Asada, H. H. (2021). Inherent haptic feedback from supernumerary robotic limbs. *IEEE Trans. Haptics.* 14, 123–131. doi: 10.1109/TOH.2020.3017548
- Guilbert, A. (2020). About the existence of a horizontal mental pitch line in non-musicians. *Laterality.* 25, 215–228. doi: 10.1080/1357650X.2019.1646756
- Hamilton-Fletcher, G., Pisanski, K., Reby, D., Stefanczyk, Ward, J., and Sorokowska, A. (2017). The role of visual experience in the emergence of cross-modal correspondences. *Cognition* 175, 114–121. doi: 10.1016/j.cognition.2018.02.023
- Hamilton-Fletcher, G., Wright, T. D., and Ward, J. (2016). Cross-modal correspondences enhance performance on a colour-to-sound sensory substitution device. *Multisens. Res.* 29, 337–363. doi: 10.1163/22134808-00002519
- Hernandez, J. W., Haget, A., Bleuler, H., Billard, A., and Bouri, M. (2019). Four-arm manipulation via feet interfaces. *arXiv [Preprint]* doi: 10.48550/arXiv.1909.04993
- Ho, C., and Spence, C. (2008). *The multisensory driver: Implications for ergonomic car interface design*. Farnham: Ashgate Publishing, Ltd.
- Hussain, I., Meli, L., Pacchierotti, C., Salvietti, G., and Prattichizzo, D. (2015). "Vibrotactile haptic feedback for intuitive control of robotic extra fingers," in *IEEE World Haptics Conference, WHC 2015*, Evanston, IL, 394–399. doi: 10.1109/WHC.2015.7177744
- Ihalainen, R., Kotsaridis, G., Vivas, A. B., and Paraskevopoulos, E. (2023). The relationship between musical training and the processing of audiovisual correspondences: Evidence from a reaction time task. *PLoS One* 18:e0282691. doi: 10.1371/journal.pone.0282691
- Imai, M., Kita, S., Nagumo, M., and Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition* 109, 54–65. doi: 10.1016/j.cognition.2008.07.015
- Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., et al. (2015). Sound symbolism facilitates word learning in 14-month-olds. *PLoS One* 10:e0116494. doi: 10.1371/journal.pone.0116494
- Klapetek, A., Ngo, M. K., and Spence, C. (2012). Does crossmodal correspondence modulate the facilitatory effect of auditory cues on visual search? *Atten. Percept. Psychophys.* 74, 1154–1167. doi: 10.3758/s13414-012-0317-9
- Köhler, W. (1929). An old pseudoproblem. *Naturwissenschaften* 17, 395–401.
- Körding, K. P., and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature* 427, 244–247. doi: 10.1038/nature02169
- Küssner, M. B., and Leech-Wilkinson, D. (2014). Investigating the influence of musical training on cross-modal correspondences and sensorimotor skills in a real-time drawing paradigm. *Psychol. Music* 42, 448–469. doi: 10.1177/0305735613482022
- Lin, A., Scheller, M., and Feng, F. (2021). "Feeling colours: Crossmodal correspondences between tangible 3D objects, colours and emotions," in *Proceedings of the Conference on human factors in computing systems*, Yokohama, doi: 10.1145/3411764.3445373
- Lorenzoni, V., Van den Berghe, P., Maes, P. J., De Bie, T., De Clercq, D., and Leman, M. (2019). Design and validation of an auditory biofeedback system for modification of running parameters. *J. Multim. User Interf.* 13, 167–180. doi: 10.1007/s12193-018-0283-1
- Ludwig, V. U., Adachi, I., and Matsuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20661–20665. doi: 10.1073/pnas.1112605108
- Maeda, F., Kanai, R., and Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Curr. Biol.* 14, R990–R991. doi: 10.1016/j.cub.2004.11.018
- Maes, P. J., Buhmann, J., and Leman, M. (2016). 3Mo: A model for music-based biofeedback. *Front. Neurosci.* 10:548. doi: 10.3389/fnins.2016.00548
- Makin, T. R., de Vignemont, F., and Faisal, A. A. (2017). Neurocognitive barriers to the embodiment of technology. *Nat. Biomed. Eng.* 1:14. doi: 10.1038/s41551-016-0014
- Marasco, P. D., Hebert, J. S., Sensinger, J. W., Shell, C. E., Schofield, J. S., Thumser, Z. C., et al. (2018). Illusory movement perception improves motor control for prosthetic hands. *Sci. Transl. Med.* 10:ea06990. doi: 10.1126/scitranslmed.a06990

- Marasco, P. D., Kim, K., Colgate, J. E., Peshkin, M. A., and Kuiken, T. A. (2011). Robotic touch shifts perception of embodiment to a prosthesis in targeted reinnervation amputees. *Brain* 134(Pt 3), 747–758. doi: 10.1093/brain/awq361
- Marks, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 384–394. doi: 10.1037//0096-1523.13.3.384
- Miall, R. C., Afanasyeva, D., Cole, J. D., and Mason, P. (2021). The role of somatosensation in automatic visuo-motor control: A comparison of congenital and acquired sensory loss. *Exp. Brain Res.* 239, 2043–2061. doi: 10.1007/s00221-021-06110-y
- Miall, R. C., Rosenthal, O., Ørstavik, K., Cole, J. D., and Sarlegna, F. R. (2019). Loss of haptic feedback impairs control of hand posture: A study in chronically deafferented individuals when grasping and lifting objects. *Exp. Brain Res.* 237, 2167–2184. doi: 10.1007/s00221-019-05583-2
- Moseley, G. L., Gallace, A., and Spence, C. (2012). Bodily illusions in health and disease: Physiological and clinical perspectives and the concept of a cortical 'body matrix'. *Neurosci. Biobehav. Rev.* 36, 34–46. doi: 10.1016/j.neubiorev.2011.03.013
- Mudd, S. A. (1963). Spatial stereotypes of four dimensions of pure tone. *J. Exp. Psychol.* 66, 347–352. doi: 10.1037/h0040045
- Nava, E., Grassi, M., and Turati, C. (2016). Audio-visual, visuo-tactile and audio-tactile correspondences in preschoolers. *Multisens. Res.* 29, 93–111. doi: 10.1163/22134808-00002493
- Noccaro, A. L., Raiano, M., Pinardi, D., Formica, G., and Pino, D. (2020). "A novel proprioceptive feedback system for supernumerary robotic limb," in *Proceedings of the IEEE RAS and EMBS international conference on Biomedical robotics and biomechanics*, 2020-Novem, New York, NY, 1024–1029. doi: 10.1109/BioRob49111.2020.9224450
- North, A. C. (2012). The effect of background music on the taste of wine. *Br. J. Psychol.* 103, 293–301. doi: 10.1111/j.2044-8295.2011.02072.x
- O'Brien, B., Hardouin, R., Rao, G., Bertin, D., and Bourdin, C. (2020). Online sonification improves cycling performance through kinematic and muscular reorganisations. *Sci. Rep.* 10, 20929. doi: 10.1038/s41598-020-76498-0
- Orchard-Mills, E., Alais, D., and Van der Burg, E. (2013). Cross-modal associations between vision, touch, and audition influence visual search through top-down attention, not bottom-up capture. *Atten. Percept. Psychophys.* 75, 1892–1905. doi: 10.3758/s13414-013-0535-9
- Page, D. M., George, J. A., Kluger, D. T., Duncan, C., Wendelken, S., Davis, T., et al. (2018). Motor control and sensory feedback enhance prosthesis embodiment and reduce phantom pain after long-term hand amputation. *Front. Hum. Neurosci.* 12:352. doi: 10.3389/fnhum.2018.00352
- Parise, C. V., and Spence, C. (2009). "When birds of a feather flock together": Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS One* 4:e5664. doi: 10.1371/journal.pone.0005664
- Parise, C., and Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neurosci. Lett.* 442, 257–261. doi: 10.1016/j.neulet.2008.07.010
- Pavani, F., Spence, C., and Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychol. Sci.* 11, 353–359. doi: 10.1111/1467-9280.00270
- Pellegrino, G., Pinardi, M., Schuler, A. L., Kobayashi, E., Masiero, S., Marioni, G., et al. (2022). Stimulation with acoustic white noise enhances motor excitability and sensorimotor integration. *Sci. Rep.* 12:13108. doi: 10.1038/s41598-022-17055-9
- Pinardi, M., Noccaro, A., Raiano, L., Formica, D., and Di Pino, G. (2023). Comparing end-effector position and joint angle feedback for online robotic limb tracking. *PLoS One* (in press). doi: 10.1371/journal.pone.0286566
- Pinardi, M. L., Raiano, A., Noccaro, D., Formica, G., and Pino, D. (2021). "Cartesian space feedback for real time tracking of a supernumerary robotic limb: A pilot study," in *International IEEE/EMBS conference on neural engineering, NER*, 2021-May, Rome, 889–892. doi: 10.1109/NER49283.2021.9441174
- Pinardi, M., Ferrari, F., D'Alonzo, M., Clemente, F., Raiano, L., Cipriani, C., et al. (2020a). Doublecheck: A sensory confirmation is required to own a robotic hand, sending a command to feel in charge of it. *Cogn. Neurosci.* 11, 216–228. doi: 10.1080/17588928.2020.1793751
- Pinardi, M., Raiano, L., Formica, D., and Di Pino, G. (2020b). Altered proprioceptive feedback influences movement kinematics in a lifting task. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2020, 3232–3235. doi: 10.1109/EMBC44109.2020.9176252
- Pramudya, R. C., Choudhury, D., Zou, M., and H-S, Seo (2020). "Bitter touch": Cross-modal associations between hand-feel touch and gustatory cues in the context of coffee consumption experience. *Food Qual. Preferen.* 83:103914. doi: 10.1016/j.foodqual.2020.103914
- Proske, U., and Gandevia, S. C. (2012). The proprioceptive senses: Their roles in signaling body shape, body position and movement, and muscle force. *Physiol. Rev.* 92, 1651–1697. doi: 10.1152/physrev.00048.2011
- Raspovic, S., Capogrosso, M., Petrini, F. M., Bonizzato, M., Rigosa, J., Di Pino, G., et al. (2014). Restoring natural sensory feedback in real-time bidirectional hand prostheses. *Sci. Transl. Med.* 6:222ra19. doi: 10.1126/scitranslmed.3006820
- Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., and Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition* 99, 113–129. doi: 10.1016/j.cognition.2005.01.004
- Segura Meraz, N., Shikida, H., and Hasegawa, Y. (2017). "Auricularis muscles based control interface for robotic extra thumb," in *Proceedings of the international symposium on micro-nanomechanics and human science (MHS)*, Nagoya. doi: 10.3233/RNN-150579
- Shadmehr, R., Smith, M. A., and Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* 33, 89–108. doi: 10.1146/annurev-neuro-060909-153135
- Slobodenyuk, N., Jraissati, Y., Kanso, A., Ghanem, L., and Elhadj, I. (2015). Cross-modal associations between color and haptics. *Atten. Percept. Psychophys.* 77, 1379–1395. doi: 10.3758/s13414-015-0837-1
- Sobajima, M., Sato, Y., Xufeng, W., and Hasegawa, Y. (2016). "Improvement of operability of extra robotic thumb using tactile feedback by electrical stimulation," in *2015 International Symposium on Micro-NanoMechatronics and Human Science, MHS 2015*, (Nagoya), 3–5. doi: 10.1109/MHS.2015.7438269
- Song, J. H. (2019). The role of attention in motor control and learning. *Curr Opin Psychol.* 29, 261–265. doi: 10.1016/j.copsyc.2019.08.002
- Sourav, S., Kekunnaya, R., Shareef, I., Banerjee, S., Bottari, S., and Röder, B. (2019). A protracted sensitive period regulates the development of cross-modal sound-shape associations in humans. *Psychol. Sci.* 30, 1473–1482. doi: 10.1177/0956797619866625
- Speed, L. J., Croijmans, I., Dolscheid, S., and Majid, A. (2021). Crossmodal associations with olfactory, auditory, and tactile stimuli in children and adults. *i-Perception* 12:20416695211048513. doi: 10.1177/20416695211048513
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Atten. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7
- Spence, C. (2014). The skin as a medium for sensory substitution. *Multisens. Res.* 27, 293–312. doi: 10.1163/22134808-00002452
- Spence, C. (2020). Food and beverage flavour pairing: A critical review of the literature. *Food Res. Int.* 133:109124. doi: 10.1016/j.foodres.2020.109124
- Spence, C. (2022). Exploring group differences in the crossmodal correspondences. *Multisens. Res.* 35, 495–536. doi: 10.1163/22134808-bja10079
- Spence, C., and Deroy, O. (2012). Crossmodal correspondences: Innate or learned?. *i-Perception* 3, 316–318. doi: 10.1068/i0526ic
- Spence, C., and Wang, Q. (2015). Wine and music (I): On the crossmodal matching of wine and music. *Flavour* 4, 1–14. doi: 10.1186/S13411-015-0045-X
- Sun, X., Li, X., Ji, L., Han, F., Wang, H., Liu, Y., et al. (2018). An extended research of crossmodal correspondence between color and sound in psychology and cognitive ergonomics. *PeerJ* 6:e4443. doi: 10.7717/peerj.4443
- Tong, J., Li, L., Bruns, P., and Röder, B. (2020). Crossmodal associations modulate multisensory spatial integration. *Atten. Percept. Psychophys.* 82, 3490–3506. doi: 10.3758/s13414-020-02083-2
- Umezawa, K., Suzuki, Y., Ganesh, G., and Miyawaki, Y. (2022). Bodily ownership of an independent supernumerary limb: An exploratory study. *Sci. Rep.* 12:2339. doi: 10.1038/s41598-022-06040-x
- von Hornbostel, E. M. (1931). Ueber geruchshelligkeit. *Pflügers Archiv Gesamte Physiol. Menschen Tiere* 228, 517–537. doi: 10.1007/BF01755351
- Walker, P., Bremner, J. G., Lunghi, M., Dolscheid, S., Barba, D., and Simion, F. (2018). Newborns are sensitive to the correspondence between auditory pitch and visuospatial elevation. *Dev. Psychobiol.* 60, 216–223. doi: 10.1002/dev.21603
- Walker-Andrews, A., Lewkowicz, D. J., and Lickliter, R. (1994). "Taxonomy for intermodal relations," in *The development of intersensory perception: Comparative perspectives*, eds D. J. Lewkowicz and R. Lickliter (Mahwah, NJ: Lawrence Erlbaum Associates, Inc), 39–56.
- Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882. doi: 10.1126/science.7569931
- Wright, O., Jraissati, Y., and Özçelik, D. (2017). Cross-modal associations between color and touch: Mapping haptic and tactile terms to the surface of the munsell color solid. *Multisens. Res.* 30, 691–715. doi: 10.1163/22134808-00002589
- Zhang, G., Wang, W., Qu, J., Li, H., Song, X., and Wang, Q. (2021). Perceptual influence of auditory pitch on motion speed. *J. Vis.* 21:11. doi: 10.1167/jov.21.10.11
- Zollo, L., Di Pino, G., Ciancio, A. L., Ranieri, F., Cordella, F., Gentile, C., et al. (2019). Restoring tactile sensations via neural interfaces for real-time force-and-slippage closed-loop control of bionic hands. *Sci. Robot.* 4:eaa9924. doi: 10.1126/scirobotics.aa9924